IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Appellants:	Sherif YACOUB et al.	§	Confirmation No.:	6342
		§		
Serial No.:	10/773,392	§	Group Art Unit:	2626
		§		
Filed:	02/06/2004	§	Examiner:	L. St. Cyr
		§		
For:	Automated Speech Recognition	§	Docket No.:	200310469-1
		§		

SUPPLEMENTAL APPEAL BRIEF

Mail Stop Appeal Brief – Patents   Date: August 11, 2008
Commissioner for Patents
PO Box 1450
Alexandria, VA 22313-1450

Sir:

Please replace the Appeal Brief filed earlier today with this Supplemental Appeal Brief. A Notice of Appeal was electronically filed on June 11, 2008.

TABLE OF CONTENTS

I. REAL PARTY IN INTEREST

The real party in interest is Hewlett-Packard Development Company, L.P. (HPDC), a Texas Limited Partnership, having its principal place of business in Houston, Texas. HPDC is a wholly owned affiliate of Hewlett-Packard Company (HPC). The Assignment from the inventors to HPDC was recorded on February 6, 2004, at Reel/Frame 014968/0891.

II.   RELATED APPEALS AND INTERFERENCES

Appellants are unaware of any related appeals or interferences.

III.   **STATUS OF THE CLAIMS**

Originally filed claims: 1-24.

Claim cancellations: 7, 8, 14, 19, 20, 23 and 24.

Added claims: None.

Presently pending claims: 1-6, 9-13, 15-18, 21 and 22.

Presently appealed claims: 1-6, 9-13, 15-18, 21 and 22.

IV.   **STATUS OF THE AMENDMENTS**

No claims were amended after the final Office action dated April 10, 2008.

V.   SUMMARY OF THE CLAIMED SUBJECT MATTER

The citations below to Appellants' disclosure and figures are merely illustrative of where support can be found for the claimed inventions that have been summarized. Other support may be present as well.

Automatic speech recognition (ASR) systems convert spoken words into computer-readable representations. An ASR engine may be the subject of a license. The license may specify the maximum number of simultaneous connections that may be established with the ASR engine. Unfortunately, the particular maximum number of connections may not always be enough. Appellants' contribution addresses this problem.[1]

The invention of claim 1 is a system[2] that comprises a first speech recognition engine,[3] a second speech recognition engine,[4] and evaluation logic[5] coupled to the first and second speech recognition engines. The evaluation logic evaluates the first and second speech recognition engines based on evaluation signals from a user[6] and, based in part on the evaluation, selects one of the speech recognition engines to process additional speech signals from the user.[7] The first speech recognition engine permits a plurality of ports to be used on behalf of a plurality of users.[8] The system further comprises a port monitor[9] coupled to the first speech recognition engine and to the evaluation logic. The port monitor determines a number of currently available ports.[10] If the number of

---

[1] See Appellants' disclosure page 1 line 1 of para. [0001] through line 7 of para. [0002].

[2] Fig. 2, 114. Appellants' disclosure p. 3 line 1 of para. [0013].

[3] Fig. 2, 212. Appellants' disclosure p. 3 line 4 of para. [0013].

[4] Fig. 2, 214. Appellants' disclosure p. 3 line 5 of para. [0013].

[5] Fig. 2, 210. Appellants' disclosure p. 3 line 4 of para. [0013].

[6] Appellants' disclosure p. 5 lines 7-8 of para. [0019].

[7] Appellants' disclosure p. 5 lines 10-12 of para. [0019].

[8] Appellants' disclosure p. 4 lines 1-3 of para. [0017].

[9] Fig. 2, 208. Page 3 line 4 of para. [0013].

[10] Appellants' disclosure p. 7 lines 3-5 of para. [0025].

currently available ports exceeds a threshold, the port monitor causes the first speech recognition engine to be selected and used.[11]

The invention of claim 9 is a system[12] that comprises first means for recognizing speech,[13] second means for recognizing speech,[14] and means for evaluating[15] a parameter associated with the first and second means for recognizing speech based on evaluation voice input from a user[16] during a session and, based on the evaluation, for selecting one of the first and second means for recognizing speech.[17] The system also comprises means for monitoring[18] a number of available ports[19] associated with the first means for recognizing speech and for selecting the first means for recognizing speech if the number of available ports exceeds a threshold.[20]

The invention of claim 15 is directed to a method that comprises evaluating an evaluation set of utterances from a user during a session.[21] Based on evaluating the evaluation set of utterances, the method further comprises selecting between a first speech recognition engine[22] and a second speech recognition engine[23] for the remainder of the session.[24] The method also comprises automatically selecting the first speech recognition engine if a number

---

[11] Appellants' disclosure p. 7 lines 8-10 of para. [0025].

[12] Fig. 2, speech recognition module 114. Appellants' disclosure p. 3 line 1 of para. [0013].

[13] Fig. 2, primary ASR engine 212. Appellants' disclosure p. 3 line 4 of para. [0013].

[14] Fig. 2, secondary ASR engine 214. Appellants' disclosure p. 3 line 5 of para. [0013].

[15] Fig. 2, evaluator 210. Appellants' disclosure p. 3 line 4 of para. [0013].

[16] Appellants' disclosure p. 5 lines 7-8 of para. [0019].

[17] Appellants' disclosure p. 5 lines 10-12 of para. [0019].

[18] Fig. 2, port monitor 208. Page 3 line 4 of para. [0013].

[19] Appellants' disclosure p. 7 lines 3-5 of para. [0025].

[20] Appellants' disclosure p. 7 lines 8-10 of para. [0025].

[21] Appellants' disclosure p. 5 lines 7-8 of para. [0019].

[22] Fig. 2, 212. Appellants' disclosure p. 3 line 4 of para. [0013].

[23] Fig. 2, 214. Appellants' disclosure p. 3 line 5 of para. [0013].

[24] Appellants' disclosure p. 5 lines 10-12 of para. [0019].

of available ports associated with the first speech recognition[25] engine exceeds a predetermined value.[26]

Claim 21 is directed to a storage medium containing code[27] that can be loaded into a computer[28] and executed by a processor in the computer. The code causes the computer to evaluate an evaluation set of utterances from a user,[29] and based on the evaluation of the evaluation set of utterances, select between a first speech recognition engine and a second speech recognition engine.[30] The code also causes the computer to determine a number of available ports associated with the first speech recognition engine[31] and to automatically select the first speech recognition engine if the number of available ports is above a threshold.[32]

---

[25] Appellants' disclosure p. 7 lines 3-5 of para. [0025].

[26] Appellants' disclosure p. 7 lines 8-10 of para. [0025].

[27] Fig. 1, 110. Appellants' disclosure pp. 3-5 lines 5-9 of para. [0010].

[28] Fig. 1, 102. Appellants' disclosure p. 3 line 3 of para. [0010].

[29] Appellants' disclosure p. 5 lines 7-8 of para. [0019].

[30] Appellants' disclosure p. 5 lines 10-12 of para. [0019].

[31] Appellants' disclosure p. 7 lines 3-5 of para. [0025].

[32] Appellants' disclosure p. 7 lines 8-10 of para. [0025].

VI.   GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

Whether claims 1-4, 6, 9-13, 15-18, 21 and 22 are obvious over Endo (U.S. Pat. No. 7,228,275) in view of Johnson (U.S. Pat. No. 6,728,671) under 35 U.S.C. § 103(a).

Whether claim 5 is obvious over Endo in view of Johnson and Kemble (U.S. Pat. No. 7,072,837) under 35 U.S.C. § 103(a).

VII.   ARGUMENT

A.   Overview of Endo

Endo teaches a speech recognition system with multiple speech recognizers. See Endo, Title. Endo states that "[e]ach of the first and second speech recognizers outputs first and second recognized speech texts and first and second associated confidence scores, respectively, and the decision module selects either the first or the second speech text depending upon which of the first or second confidence score is higher." Endo, Abstract.

B.   Overview of Johnson

Johnson is directed to a speech recognition system having a fixed number of speech recognition channels. Rather than dropping incoming calls or putting customers on hold when all current channels are otherwise used, Johnson teaches incorporating time delays into the output play mode so that a given speech recognition channel can be used for a second caller while a first caller is listening to message that has been drawn out (lengthened) with the delay. The delay mode enables a single speech recognition channel to be shared among multiple users, each one taking turns using the channel. The delays make channel sharing possible. See Johnson col. 1, line 55 through col. 2, line 58; col. 8, line 29 through col. 9, line 2; col. 12, line 65 through col. 13, line 4. Johnson compares channel usage to a threshold to determine when to transition to the delay modes. The decision point in the Johnson system is when to implement delays during message playback, not whether to switch from one speech recognition engine to another.

C.   Rejection of claims 1-4, 6, 9-13, 15-18, 21 and 22

Independent claim 1 requires, among other things,

> wherein the first speech recognition engine permits a plurality of ports to be used on behalf of a plurality of users and the system further comprises a port monitor coupled to the first speech recognition engine and to the evaluation logic, wherein the port monitor determines a number of currently available ports and, if the number of currently available ports exceeds a threshold, causes the first speech recognition engine to be selected and used.

Appellants respectfully submit that this claim element is not taught or even suggested by either Endo or Johnson.

As explained above, Endo teaches a speech recognition system wherein "[e]ach of the first and second speech recognizers outputs first and second recognized speech texts and first and second associated confidence scores, respectively, and the decision module selects either the first or the second speech text depending upon which of the first or second confidence score is higher." Endo, Abstract. As the Examiner seems to concede, no mention is made anywhere within Endo of monitoring utilization levels of any of the speech recognizers, nor of switching between speech recognizers based, even in part, upon the level of utilization of a recognizer.

Johnson teaches "first, determining a usage level of the plurality of ASR (automatic speech recognition) input channels; and second, when the usage level of the plurality of ASR input channels is greater than a first predetermined threshold, providing an associated delay mode for a message output on one or more output channels of the plurality of output channels." Johnson, col. 12 lines 65-67 through col. 13, lines 1-4. The delay mode is used during playback of a message, not by an ASR. No mention is made anywhere within Johnson of switching between ASRs based, even in part, upon the level of utilization of the ASR. Johnson switches into and out of a delay mode based on channel utilization, which is not at all the same as switching between speech recognition channels.

In response to this argument, the Examiner stated that "Johnson teaches an automatic speech recognition component having a plurality of ASR input channels; when the number of ASR input channels in use was not greater than the first predetermined threshold, the method may return or stay in no delay mode." Final Office Action at page 2. While Johnson may indeed have such a teaching, the claim requires something very different. Claim 1 requires causing a particular speech recognition engine to be selected and used based on the number of available ports exceeding a threshold. Johnson concerns, as the

Examiner has observed, using a delay mode, not a particular speech recognition engine, based on the level of ASR input channel usage.

Thus, neither Endo nor Johnson teaches or even suggests a port monitor that "if the number of currently available ports exceeds a threshold, causes the first speech recognition engine to be selected and used," as required by independent claim 1. Incorporating Johnson' teaching of delay mode selection to delay an outgoing message would not at all solve the problem faced by Appellants. Thus, one of ordinary skill in the art would not have been motivated to seek out and use Johnson's teaching. Further none of the cited art, either alone or together, overcomes the deficiencies of Endo and Johnson. For at least these reasons, Appellants respectfully submit that the Examiner erred in rejecting claim 1 and dependent claims 2, 4, and 6 over Endo in view of Johnson.

Independent claims 9, 15 and 21, which were rejected by the Examiner for the same reasons as claim 1, include limitations similar to those of independent claim 1. Thus, for at least the same reasons as those presented with regard to claim 1, Appellants respectfully submit that the Examiner erred in rejecting independent claims 9, 15 and 21, as well as those claims that respectively depend upon them.

D.   Rejection of claim 5

Regarding the rejection of dependent claim 5 as allegedly obvious over Endo in view of Johnson and further in view of Kemble, Appellants respectfully note that, because claim 5 includes all of the limitations of independent claim 1, and because Kemble fails to overcome the deficiencies of Endo and Johnson discussed above, the Examiner erred in rejecting dependent claim 5 for at least the same reason as for claim 1.

E.   Conclusion

For the reasons stated above, Appellants respectfully submit that the Examiner erred in rejecting all pending claims. It is believed that no extensions of time or fees are required, beyond those that may otherwise be provided for in documents accompanying this paper. However, in the event that additional extensions of time are necessary to allow consideration of this paper, such

extensions are hereby petitioned under 37 C.F.R. § 1.136(a), and any fees required (including fees for net addition of claims) are hereby authorized to be charged to Hewlett-Packard Development Company's Deposit Account No. 08-2025.

Respectfully submitted,

/Jonathan M. Harris/

Jonathan M. Harris
PTO Reg. No. 44,144
CONLEY ROSE, P.C.
(713) 238-8000 (Phone)
(713) 238-8008 (Fax)
ATTORNEY FOR APPELLANTS

HEWLETT-PACKARD COMPANY
Intellectual Property Administration
Legal Dept., M/S 35
P.O. Box 272400
Fort Collins, CO 80527-2400

VIII. CLAIMS APPENDIX

1. (Previously presented) A system, comprising:

a first speech recognition engine;

a second speech recognition engine; and

evaluation logic coupled to the first and second speech recognition engines, the evaluation logic evaluates the first and second speech recognition engines based on evaluation signals from a user and, based in part on the evaluation, selects one of said speech recognition engines to process additional speech signals from the user;

wherein the first speech recognition engine permits a plurality of ports to be used on behalf of a plurality of users and the system further comprises a port monitor coupled to the first speech recognition engine and to the evaluation logic, wherein the port monitor determines a number of currently available ports and, if the number of currently available ports exceeds a threshold, causes the first speech recognition engine to be selected and used.

2. (Previously presented) The system of claim 1 further comprising a switch coupled to the first and second speech recognition engines and the evaluation logic, wherein, based on the evaluation, the evaluation logic causes the switch to release a connection to the speech recognition engine that was not selected.

3. (Original) The system of claim 1 further comprising a communications mechanism and, based on the evaluation, the evaluation logic selects the communications mechanism that is not the first or second speech recognition engines.

4. (Previously presented) The system of claim 1 wherein if the number of currently available ports does not exceed the threshold, the evaluation logic compares outputs from the first and second speech recognition engines and selects the second speech recognition engine if the outputs are identical.

5.   (Previously presented) The system of claim 1 wherein if the number of currently available ports does not exceed the threshold, the evaluation logic determines a response time for each of the first and second speech recognition engines and selects the second speech recognition engine if the response time of the second speech recognition engine is equal to or shorter than the response time of the first speech recognition engine.

6.   (Previously presented) The system of claim 1 wherein if the number of currently available ports does not exceed the threshold, the evaluation logic receives a first confidence score from the first speech recognition engine and a second confidence score from the second speech recognition engine and selects the second speech recognition engine if the confidence score of the second speech recognition engine is equal to or higher than a threshold.

7.-8.   (Cancelled).

9.   (Previously presented) A system, comprising:
        first means for recognizing speech;
        second means for recognizing speech;
        means for evaluating a parameter associated with the first and second
                means for recognizing speech based on evaluation voice input from
                a user during a session and, based on the evaluation, for selecting
                one of said first and second means for recognizing speech; and
        means for monitoring a number of available ports associated with the first
                means for recognizing speech and for selecting the first means for
                recognizing speech if the number of available ports exceeds a
                threshold.

10.   (Original) The system of claim 9 further comprising means for releasing the first or second means for recognizing speech that is not selected.

11.   (Original) The system of claim 9 wherein the means for evaluating a parameter comprises means for assessing the relative accuracy of the first and second means for recognizing speech.

12.   (Original) The system of claim 9 wherein the means for evaluating a parameter comprises means for assessing the relative performance of the first and second means for recognizing speech.

13.   (Original) The system of claim 9 wherein the first and second means for recognizing speech comprise a means for determining a confidence score associated with the voice input.

14.   (Cancelled).

15.   (Previously presented) A method, comprising:
        evaluating an evaluation set of utterances from a user during a session;
        based on evaluating the evaluation set of utterances, selecting between a
            first speech recognition engine and a second speech recognition
            engine for the remainder of the session; and
        automatically selecting the first speech recognition engine if a number of
            available ports associated with the first speech recognition engine
            exceeds a predetermined value.

16.   (Previously presented) The method of claim 15 wherein evaluating the evaluation set of utterances comprises determining a relative accuracy of the first and second speech recognition engines if the number of available ports associated with the first speech recognition engine falls below a predetermined value.

17.   (Previously presented) The method of claim 15 wherein evaluating the evaluation set of utterances comprises determining a relative performance of the

first and second speech recognition engines if the number of available ports associated with the first speech recognition engine falls below a predetermined value.

18.   (Previously presented) The method of claim 15 wherein evaluating the evaluation set of utterances comprises comparing a first confidence score generated by the first speech recognition engine with a second confidence score generated by the second speech recognition engine if the number of available ports associated with the first speech recognition engine falls below a predetermined value.

19.-20. (Cancelled).

21.   (Previously presented) A storage medium containing code that can be loaded into a computer and executed by a processor in the computer, the code causing the computer to:

      evaluate an evaluation set of utterances from a user;

      based on the evaluation of the evaluation set of utterances, select between a first speech recognition engine and a second speech recognition engine; and

      determine a number of available ports associated with the first speech recognition engine and to automatically select the first speech recognition engine if the number of available ports is above a threshold.

22.   (Previously presented) The storage medium of claim 21 wherein the code causes the processor to evaluate the evaluation set of utterances, if the number of available ports is below the threshold, by performing an action selected from the group consisting of comparing a relative accuracy of the first and second speech recognition engines, comparing the a relative performance of the first and second speech recognition engines, and comparing a confidence score

generated by the first and second speech recognition engines, and a combination thereof.

23.-24. (Cancelled).

IX.   EVIDENCE APPENDIX

None.

X.   RELATED PROCEEDINGS APPENDIX

None.